Final Scientific Report

Cover Page

BARD Project Number: IS-4067-07

Date of Submission of the report: August 28, 2011

Project Title: Genomic Approaches for Understanding Virulence and Resistance in the Sunflower-*Orobanche* Host-Parasite Interaction

Investigators	**Institutions**
Principal Investigator (PI): Daniel M. Joel	ARO
Co-Principal Investigator (Co-PI): Steven J. Knapp	University of Georgia
Collaborating Investigators: Ya'akov Tadmor	ARO

Keywords

Orobanche cumana;

Abbreviations commonly used in the report, in alphabetical order**:**

Appendix G6a

ESTs - Expressed sequence tags; SNP - single nucleotide polymorphism; SSRs - simple-sequence repeats.

Budget: IS: $ 157,000 US: $ 148,000 Total: $ 305,000

_____ _____

Signature
Principal Investigator

Signature
Authorizing Official, Principal Institution

Final Scientific Report

Publication Summary (numbers)

	Joint IS/US authorship	US Authors only	Israeli Authors only	Total
Refereed (published, in press, accepted) BARD support acknowledged		1	1	2
Submitted, in review, in preparation	3			3
Invited review papers				
Book chapters				
Books				
Master theses				
Ph.D. theses				
Abstracts				
Not refereed (proceedings, reports, etc.)	1			1

Postdoctoral Training: List the names and social security/identity numbers of all postdocs who received more than 50% of their funding by the grant. None

Cooperation Summary (numbers)

	From US to Israel	From Israel to US	Together, elsewhere	Total
Short Visits & Meetings		2		2
Longer Visits (Sabbaticals)				

Description Cooperation:

The Israeli team developed the hybrids and F2 families of *Orobanche cumana* populations. F_2 sunflower hybrids seeds were developed in UGA with *Shemesh* sunflower

seeds from Israel. Phenotyping of both sunflower and *Orobanche* hybrid families was conducted at the ARO. Frozen and lyophylized tissue samples, as well as DNA of both *O. cumana* and sunflower populations and hybrids were sent from the ARO to UGA for genomic analyses.

The PI visited the US and met the co-PI twice, once before submission of the research proposal, and once after its approval, for discussion and coordination of the work done by the two groups. During the first half of the project we had routine coordination by email and 'Skype meetings', as well as exchange of biological material. The co-PI left the UGA in 2009, and therefore coordination was difficult during the second part of the project. The PI invited the UGA post-doc to the International Congress on Parasitic Plants where they met and exchanged information in preparation of joint publications.

Patent Summary: None

Abstract

- **Oroginal Objectives:** (i) identify DNA markers linked to the avirulence (*Avr*) locus and locate the *Avr* locus through genetic mapping with an inter-race *Orobanche cumana* population; (ii) develop high-throughput fingerprint DNA markers for genotyping *O. cumana* races; (iii) identify nucleotide binding domain leucine rich repeat (NB-LRR) genes encoding *R* proteins conferring resistance to *O. cumana* in sunflower; (iv) increase the resolution of the chromosomal segment harboring Or_5 and related *R* genes through genetic and physical mapping in previously and newly developed mapping populations of sunflower; and (v) develop high-throughput DNA markers for rapidly and efficiently identifying and transferring sunflower *R* genes through marker-assisted selection.
- **Revisions made during the course of project:** Following changes in *O. cumana* race distribution in Israel, the newly arrived virulent race H was chosen for further analysis. *HA412-HO*, which was primarily chosen as a susceptible sunflower cultivar, was more resistant to the new parasite populations than var. *Shemesh*, thus we shifted sunflower research into analyzing the resistance of *HA412-HO*. We exceeded the deliverables for Objectives #3-5 by securing funding for complete physical and high-density genetic mapping of the sunflower genome, in addition to producing a complete draft sequence of the sunflower genome. We discovered limited diversity between the parents of the *O. cumana* population developed for the mapping study. Hence, the developed DNA marker resources were insufficient to support genetic map construction. This objective was beyond the scale and scope of the funding. This objective is challenging enough to be the entire focus of follow up studies.
- **Background to the topic:** *O. cumana*, an obligate parasitic weed, is one of the most economically important and damaging diseases of sunflower, causes significant yield losses in susceptible genotypes, and threatens production in Israel and many other countries. Breeding for resistance has been crucial for protecting sunflower from *O. cumana*, and problematic because new races of the pathogen continually emerge, necessitating discovery

and deployment of new *R* genes. The process is challenging because of the uncertainty in identifying races in a genetically diverse parasite.

- **Major conclusions, solutions, achievements:** We developed a small collection of SSR markers for genetic mapping in *O. cumana* and completed a diversity study to lay the ground for objective #1. Because DNA sequencing and SNP genotyping technology dramatically advanced during the course of the study, we recommend shifting future work to SNP discovery and mapping using array-based approaches, instead of SSR markers. We completed a pilot study using a 96-SNP array, but it was not large enough to support genetic mapping in *O.cumana*. The development of further SNPs was beyond the scope of the grant. However, the collection of SSR markers was ideal for genetic diversity analysis, which indicated that *O. cumana* populations in Israel considerably differ from populations in other Mediterranean countries.
  We supplied physical and genetic mapping resources for identifying *R*-genes in sunflower responsible for resistance to *O. cumana*. Several thousand mapped SNP markers and a complete draft of the sunflower genome sequence are powerful tools for identifying additional candidate genes and understanding the genomic architecture of *O. cumana*-resistance and disease-resistance genes.
- **Implications:** The *Orobanche* SSR markers have utility in sunflower breeding and genetics programs, as well as a tool for understanding the heterogeneity of races in the field and for geographically mapping of pathotypes. The segregating populations of both *Orobanche* and sunflower hybrids are now available for QTL analyses.

Achievements

Objective #1. To study the genetic basis of virulence in *O.cumana* we selected the newly arrived virulent population (H), and a less virulent population (A), crossed them and developed F3 seed families, the virulence of which was tested in vitro (appendix Fig. 1). Altogether 79 F3 families were analyzed (appendix Fig. 2A), and their majority showed a phenotype similar to the less virulent population, which can only be explained by the inbreeding nature of this species. For genetic mapping and other applications in *O. cumana* we developed 3,000 SSRs and identified 553 SNPs (See Objective #2) and, from preliminary screening of the parents of the *O. cumana* mapping population, estimate that approximately one-third of the SSRs and half of the SNPs are polymorphic. This number should be sufficient to support genetic mapping in *O. cumana*. Because of the state of the technology, we strongly discourage investing in SSR marker development and screening, and encourage investing in SNP marker development and screening.

Objective #2. This objective was completed. Stalk DNA and tissue samples were harvested at the ARO from two greenhouse-grown Israeli *O. cumana* races (A and H) identified by screening 10 *O. cumana* populations from Israel with 20 Israeli and US sunflower host differentials. Frozen *O. cumana* samples were shipped from Israel to the US, and RNA and genomic DNA isolated from both races. DNA markers had not been previously developed for *O. cumana* races (see Objective #1). Using next-generation sequencing technologies, DNA and cDNA sequences were screened for simple sequence repeats (SSRs) and single nucleotide polymorphisms (SNPs) to support SSR and SNP marker development. We produced 450,000 454-Titanium (454-Ti) ESTs reads for *O. cumana* split between races A and H. These ESTs reads were analyzed and assembled using MIRA and other bioinformatic tools to develop an EST database for *O. cumana*. The transcript assembly was mined for SNPs and SSRs leading to the identification of 4,973 SSRs (2,865 dinucleotides, 369 trinucleotides, 502 tetra nucleotides, 185 pentanucleotides, and 133 hexanucleotides). From these SSR we selected and developed 50 EST-SSR markers for biodiversity and phylogeographic analyses and quantifying the distribution of allelic diversity within and among *O. cumana* ecotypes (natural populations). These were utilized to screen for SSR length polymorphisms among 32 DNA samples isolated from *O. cumana* populations collected in Israel and southern Europe: 31 of the 50 SSR markers produced unambiguous genotypes. Of these, 20 were

selected for future biodiversity analyses of *O. cumana*. The Israeli populations were found to significantly differ from the European (appendix Fig. 3). In addition we also found in *O. cumana* seeds originating from various field populations a different germination response to the application of different dilutions of root exudates of sunflower (appendix Fig 4), indicating that at least part of the field diversity is based on differences in the response to germination stimulants.

Objective #3. Soon after this grant was awarded, the US collaborator was awarded funding by Genome Canada (as a co-principal investigator) to sequence the sunflower genome. We achieved all of their milestones and are close to having an initial draft of the sunflower genome sequence. The existence of this work and sequence information have supplied a complete catalog of the candidate NBS-LRR genes in the region on chromosome 3 flanking *Or*$_5$. DNA markers (RGC172 and RGC181) were previously developed for two NBS-LRR encoding resistance gene candidates (RGCs) identified by screening a sunflower BAC library with overgo probes. RGC172 and RGC181 were tightly linked to *Or*$_5$, one of several *O. cumana* resistance (*R*) genes found near the upper end of linkage group 3 in sunflower. We designed an overgo probe for RGC172 which was used to isolate several BAC clones. The BAC clones were fingerprinted, assembled into contigs, and end-sequenced. The BACs were positive for RGC172 and assembled into a single contig (RGC172-C1). Single strand conformational polymorphism (SSCP) markers were developed from several BAC end sequences (BESs). Two BES-SSCP markers were genotyped in the PHC x PHD recombinant inbred line (RIL) mapping population and both co-segregated with RGC172, the DNA markers originally used to isolate the BACs. One BAC in the contig was shotgun sequenced, assembled, annotated, and analyzed. One full-length NBS-LRR encoding gene was identified. This process was repeated for a DNA marker (CRT392) located near *Or*$_5$ and the telomere.

Objective #4. Since this work was funded, the US collaborator was awarded funding (as the principal investigator) to sequence develop a 10,000 SNP array for sunflower, screen diverse germplasm, including the parents of the sunflower mapping populations developed for our BARD collaboration, and to complete the genetic mapping of as many loci as possible on the array. This work concluded last winter and resulted in genetic mapping of 8,000 SNP loci. These supply a dense catalog of mapped SNP markers for application in the BARD populations, which employed an RIL population segregating for

Or$_5$ (PHC x PHD), a new F_2-F_3 population we developed from a cross between HA412-HO and *Shemesh*, and two reference populations (RHA280 x RHA801 and NMS373 x Hopi). HA412-HO x *Shemesh*, NMS373 x *Shemesh*, and RHA801 x *Shemesh* F_1 seeds were produced by hand emasculation. F_2 populations were produced in UGA by selfing F1 plants, and the F2 seeds were shipped to Israel, where F2 plants were grown (appendix Fig. 2B) and the resulting F3 plants were phenotyped for resistance to *O. cumana* (appendix Fig. 4C). The continuous phenotypic segregation within and between F3 families indicated that the resistance is multigenic. DNA was extracted from tissue samples (appendix Fig. 4D), and the first batch of DNA samples was shipped from the ARO to UGA for DNA marker genotyping and genetic mapping. Due to limited resources the genotyping could not be done. The analysis of this biological material is waiting for further funding.

Objective #5. This objective was achieved. The previously described SNP marker discovery, development, and mapping work supplied academia and the private sector in the US, Israel, and elsewhere with 8,000 mapped SNP markers, including several hundred SNP markers in the upper arm of sunflower chromosome 3 flanking the *Or*$_5$ locus.

- **Significance of main scientific achievements or innovations.**

  The most significant finding is that *O. cumana* populations in Israel differ from the known European population for which sunflower resistances have so far been mapped. We developed for the first time SSR and SNP markers for *O. cumana*, which should be sufficient to support genetic mapping in *O. cumana*, and have utility for understanding the heterogeneity of races in the field.

  For sunflower, the SNP marker discovery development and mapping supplied 8,000 mapped SNP markers, including several hundred SNP markers in the upper arm of chromosome 3 flanking the *Or*$_5$ locus.

- **Agricultural and/or economic impacts of the research findings, if known.**

  The *O. cumana* markers have utility in sunflower breeding and genetics programs, and as a tool for geographically mapping of the parasite pathotypes. The segregating populations of both *O. cumana* and sunflower hybrids are now available for QTL analyses.

.

Details of cooperation: The Israeli team developed the hybrids and F2 families of *Orobanche cumana* populations and prepared the sunflower F2 and F3 plants. F_2 sunflower hybrid seeds were developed in UGA with *Shemesh* sunflower seeds from Israel. Phenotyping of both sunflower and *Orobanche* hybrid families was conducted at the ARO. Frozen and lyophylized tissue samples, as well as DNA of both *O. cumana* UGA team conducted the genomic analyses of this genetic material. D. Joel visited the US and met S. Knapp twice, once before submission of the research proposal, and once after its approval, for discussion and coordination of the work done by the two groups. During the first half of the project we had routine coordination by email and 'Skype meetings', as well as exchange of biological material. The American co-PI left the UGA in 2009for Monsanto before completion of the project, and consequently communication and coordination were difficult during the second part of the project. The PI invited the UGA post-doc to the International Congress on Parasitic Plants where they met and exchanged information in preparation of joint publications on *O. cumana*.

Publications:

Joel D.M., Chaudhuri S.K., Plakhine, D., Ziadne, H. and Steffens J.C. (2011) Dehydrocostus lactone is exuded from sunflower roots and stimulates germination of the root parasite *Orobanche cumana*. *Phytochemistry* 72: 624–634.

Kane NC, Gill N, K, King MG, Bowers JE, Berges H, Gouzy J, Bachlava E, Langlade Z, Lai Z, Stewart M, Burke JM, Vincourt P, Knapp SJ, Rieseberg LH (2011) Progress towards a reference genome for sunflower. *Botany* (in press)

Breton C, Tadmor Y, Joel D, Knapp S. (2010) Genomic approach using the pyrosequensing method to unravel genetic diversity of Broomrape (*Orobanche cumana* Wallr). Abstract. *Evolution 2010*, Portland State University, Oregon USA.

In preparation:

Bachlava EB, Taylor C, Bowers J, Burke JM, Knapp SJ. SNP Discovery and Validation in Sunflower.

Bowers J, Bachlava EB, Taylor C, Barb J, Michelmore RW, Burke JM, Knapp SJ. Genetic Mapping of 8,000 SNPs in the Sunflower Genome.

Acknowledgement:

We thank Dr. Catherine Breton for her efforts to complete the *Orobanche* parts of the project and for her continuous work on the project after returning to France, which helped in the preparation of this report. We thank Dr. Dina Plakhine and Mr. Hammam Ziadne for their dedication and excellent work in the development of the *O. cumana* hybrids and in phenotyping both sunflower and *O. cumana*. We also thank Kibbutz Sha'ar HaAmakin and especially Mr. Eitan Herman for kindly supplying sunflower seeds for this research project. The sunflower parts of the project were also supported by funds from other sources; the sunflower SNP discovery and mapping grant was $1,000,000 and the sunflower genome sequencing grant was $9,000,000, far beyond the scope of this BARD project.

Appendix

Table of contents:

- Figures 1-4.
- Unpublished (in preparation) data briefly summarized.
- Published papers: Joel el al. 2011.
- In press paper: Kane et al. 2011.

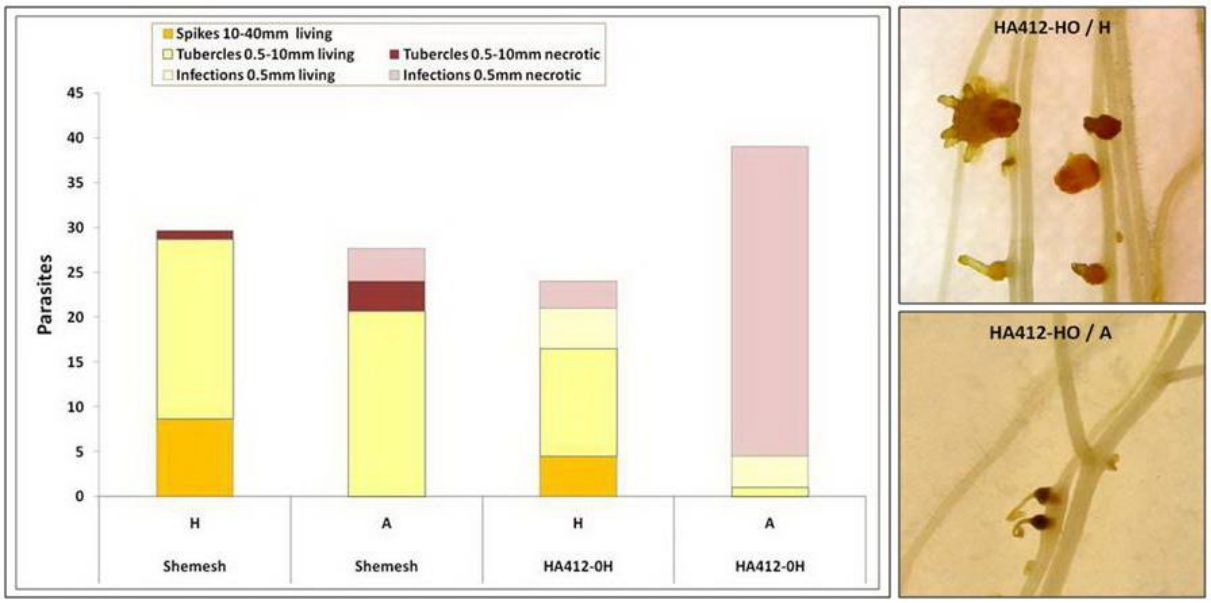

Fig1. Graph summarizing the differential attack of the two sunflower cultivars *Shemesh* and *HA412-HO* by the O. *cumana* races A and H, *and* micrographs of the differential attack of sunflower cultivar HA412-HO in polyethylene bags by two *O. cumana* populations. Root hypersensitive response - only with race A.

Fig 2. A. Phenotyping *O. cumana* hybrids on sunflower plants in the greenhouse.
B. F2 sunflower plants grown in Newe-Ya'ar Research Center.
C. Phenotyping F3 sunflower plants in polyethylene bags containing *O. cumana* seeds.
D. F3 sunflower samples before DNA extraction.

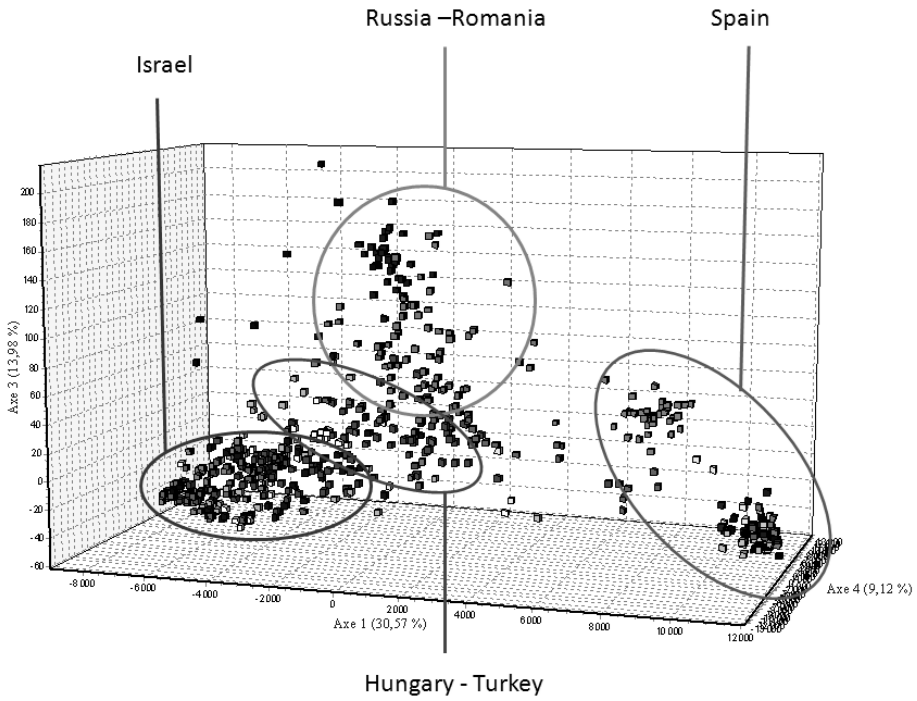

Fig 3. SSR Diversity results showing that the *O. cumana* populations in Israel are distinct from those of other Mediterranean populations.

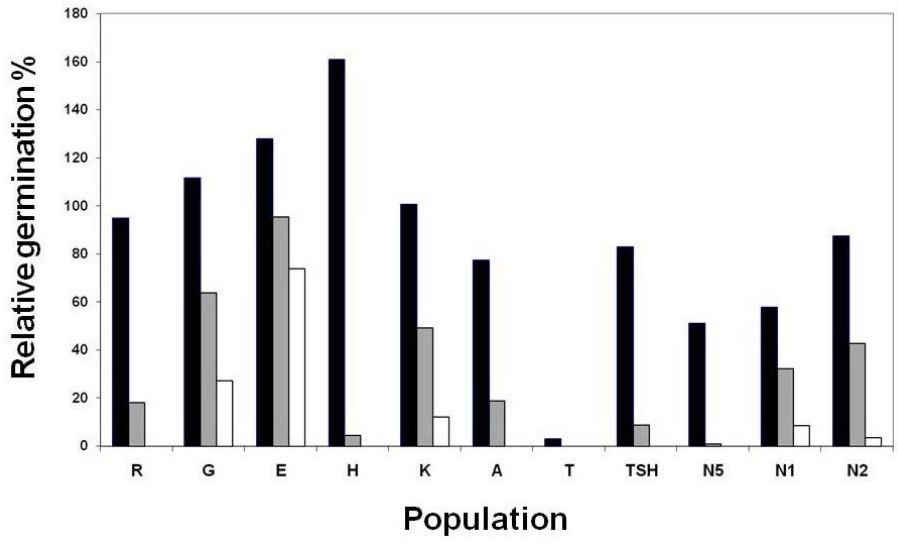

Fig 4. Germination response of seeds collected from eleven *O. cumana* populations in Israel (including populations A and H), each tested with three concentrations of root exudates collected from a sunflower plant (published in Joel et al. 2011. Phytochemistry 72: 624–634).

- Papers in preparation briefly summarized.

SNP Discovery and Validation in Sunflower

Bachlava EB, Taylor C, Bowers J, Burke JM, Knapp SJ

We describe a large-scale SNP discovery and genotyping effort to facilitate high-density genetic mapping and fine mapping of quantitative traits in sunflower. We obtained approximately 6 Gb of sunflower transcriptome using next-generation sequening technologies, developed an updated sunflower de novo long-read EST reference assembly, and mapped short-read ESTs for SNP discovery. We identified several thousands high-quality SNPs among 9 elite sunfower lines using a customized bioinformatic pipeline, and selected the optimal 10,640 SNPs for the development of a high-throughput SNP genotyping array with Illumina's iSelect Infinium platform.

Genetic Mapping of 8,000 SNPs in the Sunflower Genome

Bowers J, Bachlava EB, Taylor C, Barb J, Michelmore RW, Burke JM, Knapp SJ.

A composite genetic map of sunflower with approximately 10,000 loci was constructed. The combined map was based on 4 separate maps with from 3000-5000 loci each, representing a diverse set of cultivated and wild parents. The maps were based on an Illumina chip corresponding to 9480 SNPs and integrated with over 1500 previously mapped SSR loci.